

1601
#2
TECH CENTER 1600/2900

13 2001

RECEIVED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/880,881

DATE: 09/28/2001

TIME: 13:20:55

Input Set : N:\Crf3\RULE60\09881088.txt

Output Set: N:\CRF3\09282001\I880881.raw

SEQUENCE LISTING

C--> 5 (1) GENERAL INFORMATION:

- 7 (i) APPLICANT: Burnham, Martin K.
8 Lonetto, Michael A.
9 Warren, Patrick V.

C--> 11 (ii) TITLE OF INVENTION: NOVEL HIS5

14 (iii) NUMBER OF SEQUENCES: 7

16 (iv) CORRESPONDENCE ADDRESS:

- 17 (A) ADDRESSEE: Dechert Price & Rhoads
18 (B) STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
19 (C) CITY: Philadelphia
20 (D) STATE: PA
21 (E) COUNTRY: US
22 (F) ZIP: 19103

24 (v) COMPUTER READABLE FORM:

- 25 (A) MEDIUM TYPE: Diskette
26 (B) COMPUTER: IBM Compatible
27 (C) OPERATING SYSTEM: DOS
28 (D) SOFTWARE: FastSEQ for Windows Version 2.0

30 (vi) CURRENT APPLICATION DATA:

- C--> 31 (A) APPLICATION NUMBER: US/09/880,881
C--> 32 (B) FILING DATE: 14-Jun-2001
33 (C) CLASSIFICATION:

35 (vii) PRIOR APPLICATION DATA:

- 36 (A) APPLICATION NUMBER: 08/953,139
37 (B) FILING DATE: 1997-10-17

41 (viii) ATTORNEY/AGENT INFORMATION:

- 42 (A) NAME: Dickinson, Todd Q
43 (B) REGISTRATION NUMBER: 28,354
44 (C) REFERENCE/DOCKET NUMBER: GM10056-01

46 (ix) TELECOMMUNICATION INFORMATION:

- 47 (A) TELEPHONE: 215-994-2252
48 (B) TELEFAX: 215-994-2222
49 (C) TELEX:

52 (2) INFORMATION FOR SEQ ID NO: 1:

54 (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 579 base pairs
56 (B) TYPE: nucleic acid
57 (C) STRANDEDNESS: double
58 (D) TOPOLOGY: linear

61 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

63	ATGATTGTCA TCGTTGATTA TGGATTAGGG AATATTAGTA ATGTAAAACG CGCTATTGAA	60
64	CATTTAGGGT ATGAGGTGGT TGTCTCAAAT ACCTCAAAAA TAATCGATCA AGCAGAAACA	120
65	ATCATATTGC CCGGTGTCGG CCATTTTAAA GATGCGATGT CAGAGATAAA ACGATTAAAT	180
66	CTCAATGCAA TATTGGCTAA GAATACTGAT AAGAAGATGA TTGGTATTG TTTAGGCATG	240
67	CAATTAATGT ATGAGCATAG TGATGAAGGC GATGCATCTG GATTAGGGTT TATCCCAGGA	300
68	AATATTTTCG GTATCCAAAC AGAATACCCA GTGCCGCACT TAGGTTGGAA TAATTTAGTG	360

ENTERED

RAW SEQUENCE LISTING

DATE: 09/28/2001

PATENT APPLICATION: US/09/880,881

TIME: 13:20:55

Input Set : N:\Crf3\RULE60\09881088.txt

Output Set: N:\CRF3\09282001\I880881.raw

```

69 AGTAAGCATC CTATGTTAAA TCAAGATGTT TACTTCGTAC ATTCTTACCA AGCGCCGATG 420
70 TCAGAAAATG TCATTGCATA TCGCAGTAT GGGGCTGATA TTCCGGCAAT TGTTCAATTT 480
71 AACCAATTATA TTGGTATTCA ATTCCATCCT GAAAAAAGCG GTACATATGG GTTACAAATT 540
72 TTGCGTCAGG CAATACAAGG GGGATTTATA AATGATTGA 579
74 (2) INFORMATION FOR SEQ ID NO: 2:
76 (i) SEQUENCE CHARACTERISTICS:
77 (A) LENGTH: 192 amino acids
78 (B) TYPE: amino acid
79 (C) STRANDEDNESS: single
80 (D) TOPOLOGY: linear
83 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
85 Met Ile Val Ile Val Asp Tyr Gly Leu Gly Asn Ile Ser Asn Val Lys
86 1 5 10 15
87 Arg Ala Ile Glu His Leu Gly Tyr Glu Val Val Val Ser Asn Thr Ser
88 20 25 30
89 Lys Ile Ile Asp Gln Ala Glu Thr Ile Ile Leu Pro Gly Val Gly His
90 35 40 45
91 Phe Lys Asp Ala Met Ser Glu Ile Lys Arg Leu Asn Leu Asn Ala Ile
92 50 55 60
93 Leu Ala Lys Asn Thr Asp Lys Lys Met Ile Gly Ile Cys Leu Gly Met
94 65 70 75 80
95 Gln Leu Met Tyr Glu His Ser Asp Glu Gly Asp Ala Ser Gly Leu Gly
96 85 90 95
97 Phe Ile Pro Gly Asn Ile Ser Arg Ile Gln Thr Glu Tyr Pro Val Pro
98 100 105 110
99 His Leu Gly Trp Asn Asn Leu Val Ser Lys His Pro Met Leu Asn Gln
100 115 120 125
101 Asp Val Tyr Phe Val His Ser Tyr Gln Ala Pro Met Ser Glu Asn Val
102 130 135 140
103 Ile Ala Tyr Ala Gln Tyr Gly Ala Asp Ile Pro Ala Ile Val Gln Phe
104 145 150 155 160
105 Asn Asn Tyr Ile Gly Ile Gln Phe His Pro Glu Lys Ser Gly Thr Tyr
106 165 170 175
107 Gly Leu Gln Ile Leu Arg Gln Ala Ile Gln Gly Gly Phe Ile Asn Asp
108 180 185 190
110 (2) INFORMATION FOR SEQ ID NO: 3:
112 (i) SEQUENCE CHARACTERISTICS:
113 (A) LENGTH: 576 base pairs
114 (B) TYPE: nucleic acid
115 (C) STRANDEDNESS: double
116 (D) TOPOLOGY: linear
119 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
121 ATGATTGTCA TCGTTGATTA TGGATTAGGG AATATTAGTA ATGTAAAACG CGCTATTGAA 60
122 CATTAGGGT ATGAGGTGGT TGTCTCAAT ACCTCAAAA TAATCGATCA AGCAGAAACA 120
123 ATCATATTGC CCGGTGTCGG CCATTTTAAA GATGCGATGT CAGAGATAAA ACGATTAAAT 180
124 CTCAATGCAA TATTGGCTAA GAATACTGAT AAGAAGATGA TTGGTATTTG TTTAGGCATG 240
125 CAATTAATGT ATGAGCATAG TGATGAAGGC GATGCATCTG GATTAGGGTT TATCCCAGGA 300
126 AATATTTTCGC GTATCCAAAC AGAATACCCA GTGCCGCACT TAGGTTGGAA TAATTTAGTG 360
127 AGTAAGCATC CTATGTTAAA TCAAGATGTT TACTTCGTAC ATTCTTACCA AGCGCCGATG 420

```

RAW SEQUENCE LISTING

DATE: 09/28/2001

PATENT APPLICATION: US/09/880,881

TIME: 13:20:55

Input Set : N:\Crf3\RULE60\09881088.txt

Output Set: N:\CRF3\09282001\I880881.raw

128 TCAGAAAATG TCATTGCATA TGCGCAGTAT GGGGCTGATA TTCCGGCAAT TGTTC AATTT 480
 129 AACCAATTATA TTGGTATTCA ATTCCATCCT GAAAAAAGCG GTACATATGG GTTACAAATT 540
 130 TTGCGTCAGG CAATACAAGG GGGATTTATA AATGAT 576

132 (2) INFORMATION FOR SEQ ID NO: 4:

134 (i) SEQUENCE CHARACTERISTICS:

135 (A) LENGTH: 192 amino acids

136 (B) TYPE: amino acid

137 (C) STRANDEDNESS: single

138 (D) TOPOLOGY: linear

141 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

143 Met Ile Val Ile Val Asp Tyr Gly Leu Gly Asn Ile Ser Asn Val Lys

144 1 5 10 15

145 Arg Ala Ile Glu His Leu Gly Tyr Glu Val Val Val Ser Asn Thr Ser

146 20 25 30

147 Lys Ile Ile Asp Gln Ala Glu Thr Ile Ile Leu Pro Gly Val Gly His

148 35 40 45

149 Phe Lys Asp Ala Met Ser Glu Ile Lys Arg Leu Asn Leu Asn Ala Ile

150 50 55 60

151 Leu Ala Lys Asn Thr Asp Lys Lys Met Ile Gly Ile Cys Leu Gly Met

152 65 70 75 80

153 Gln Leu Met Tyr Glu His Ser Asp Glu Gly Asp Ala Ser Gly Leu Gly

154 85 90 95

155 Phe Ile Pro Gly Asn Ile Ser Arg Ile Gln Thr Glu Tyr Pro Val Pro

156 100 105 110

157 His Leu Gly Trp Asn Asn Leu Val Ser Lys His Pro Met Leu Asn Gln

158 115 120 125

159 Asp Val Tyr Phe Val His Ser Tyr Gln Ala Pro Met Ser Glu Asn Val

160 130 135 140

161 Ile Ala Tyr Ala Gln Tyr Gly Ala Asp Ile Pro Ala Ile Val Gln Phe

162 145 150 155 160

163 Asn Asn Tyr Ile Gly Ile Gln Phe His Pro Glu Lys Ser Gly Thr Tyr

164 165 170 175

165 Gly Leu Gln Ile Leu Arg Gln Ala Ile Gln Gly Gly Phe Ile Asn Asp

166 180 185 190

168 (2) INFORMATION FOR SEQ ID NO: 5:

170 (i) SEQUENCE CHARACTERISTICS:

171 (A) LENGTH: 20 base pairs

172 (B) TYPE: nucleic acid

173 (C) STRANDEDNESS: single

174 (D) TOPOLOGY: linear

177 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

179 ATGTAAAACG CGCTATTGAA

20

181 (2) INFORMATION FOR SEQ ID NO: 6:

183 (i) SEQUENCE CHARACTERISTICS:

184 (A) LENGTH: 19 base pairs

185 (B) TYPE: nucleic acid

186 (C) STRANDEDNESS: single

187 (D) TOPOLOGY: linear

190 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

RAW SEQUENCE LISTING

DATE: 09/28/2001

PATENT APPLICATION: US/09/880,881

TIME: 13:20:55

Input Set : N:\Crf3\RULE60\09881088.txt

Output Set: N:\CRF3\09282001\I880881.raw

```
192 CCCCTTGAT TGCCTGACG 19
194 (2) INFORMATION FOR SEQ ID NO: 7:
196 (i) SEQUENCE CHARACTERISTICS:
197 (A) LENGTH: 25 base pairs
198 (B) TYPE: nucleic acid
199 (C) STRANDEDNESS: single
200 (D) TOPOLOGY: linear
203 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
205 GCTCCTAAAA GGTTACTCCA CCGGC 25
```

VERIFICATION SUMMARY

DATE: 09/28/2001

PATENT APPLICATION: US/09/880,881

TIME: 13:20:56

Input Set : N:\Crf3\RULE60\09881088.txt

Output Set: N:\CRF3\09282001\I880881.raw

L:5 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:11 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]